様　式　Ｃ－１９、Ｆ－１９、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

平成 ２７ 年　 ６ 月 １０ 日現在

機関番号：１４３０１
研究種目：若手研究(B)
研究期間：2013～2014
課題番号：２５７３００８４

研究課題名（和文）検索性とプライバシー保護性を両立する顔画像の蓄積及び検索機構の研究

研究課題名（英文）A Study of Facial Image Storage and Retrieval System with Privacy Protection

研究代表者
森村　吉貴（Morimura, Yoshitaka）

京都大学・学術情報メディアセンター・助教

研究者番号：８０５７８２７９

交付決定額（研究期間全体）：（直接経費）　　2,800,000円

研究成果の概要（和文）：本研究では画像検索技術と暗号技術の統合により、犯罪者の顔画像の検索性と善良な市民のプライバシーの保護性を両立する顔画像の蓄積及び検索機構の構築を目指した。そのために、(1)顔画像からの適切な識別符号の生成法、(2)識別符号に基づく検索正当性の確認手続きを備えた暗号化処理及び復号処理法の２点の実現方法を明らかにし、この組み合わせにより提案機構を構築・評価した。その結果、顔画像検索におけるプライバシー保護と社会的利便性の両立についての知見を得た。

研究成果の概要（英文）：We aimed to establish the facial image storage and retrieval system to search facial images of criminal suspects and to protect privacy of good citizens, based on integration of image retrieval technology and cryptographic technology. To realize it, we proposed methods to generate effective identification code from a facial image (1) and to encrypt and decrypt data with a retrieval authorization process based on the identification code (2). We implemented and evaluated the proposed system and obtain some observation to balance privacy protection and social benefit.

研究分野：情報学

キーワード：プライバシー保護　画像検索　顔画像認識　情報セキュリティ

１．研究開始当初の背景

近年、実世界から情報を獲得するセンサ機器や、その獲得された情報を蓄積する大容量ストレージは目覚ましく普及している。特に防犯に対する需要から、公共空間には多数の監視カメラが設置され、社会全体で膨大な映像群が蓄積されつつある。画像検索技術の発展は、このような映像群から指定した画像と類似する顔画像群を取り出すことを可能にした。このような技術により「犯罪者の顔画像」を指定して検索すれば、その犯罪者が過去「いつどこに居たのか」が容易に得られる。従って、監視カメラ映像からの顔画像検索の仕組みは犯罪捜査にとって非常に強力な武器の一つとなる。しかし、ある個人が「いつどこに居たのか」というのはプライバシー情報であり、社会全体で蓄積された映像群に含まれるプライバシー情報は膨大となる。もし警察が職権を濫用して犯罪者以外の市民の顔画像を検索したり、映像蓄積装置が盗難されたりすると、重大なプライバシー上の懸念が生じる。

善良な市民のプライバシーを保護しながら監視カメラの映像を有効に利用するため、得られた映像中から人物画像を画像処理により匿名化するような研究も行われている（引用文献①）。しかし、映像の撮影時点ではそれぞれ被写体となる人物が将来犯罪を行うかどうかまではわからないため、潜在的な犯罪者の顔画像を選択的に匿名化せずに残しておくということは不可能である。従って、将来の顔画像検索能力のために得られた被写体の画像を全て残しながらも、善良な市民の顔画像のプライバシーを保護することができる技術が求められる。

２．研究の目的

そこで本研究では、犯罪者の顔画像の検索能力を残しながら善良な市民の顔画像のプライバシーを保護することを目標とし、画像検索技術と暗号技術を統合することで以下のような顔画像の蓄積及び検索機構の実現を目指す。

【画像蓄積時】
・撮影された顔画像から特徴量を抽出し、個を識別する符号を生成する
・上記の識別符号と紐づく暗号鍵を使いて顔画像を暗号化し蓄積する
【画像検索時】
・警察は検索したい犯罪者の顔画像を裁判所や役所など中立性の高い機関に提出する
・中立機関は提出された犯罪者の顔画像検索の正当性を確認する
・正当と認められれば顔画像から識別符号を生成し対応する復号鍵を警察に与える
・警察は暗号化蓄積情報に対して復号を試行してゆき、成功したものを検索結果とする

例えば犯罪捜査のための顔画像検索においてこのような機構が実現できれば、警察が得られる画像検索結果は犯罪者の顔画像と同じ識別符号を持つ顔画像のみとなり、それ以外の顔画像が持つ情報を保護することができる。

３．研究の方法

（１）顔画像からの適切な識別符号の生成法

顔画像の撮影において、同一人物を異なる場所・時間で撮影した場合、類似するがそれぞれ異なる顔画像が得られる。従って本研究の対象となる監視カメラ映像の検索においては、同一人物内の特徴量変動を吸収するような識別符号の生成法が必要となる。これを換言すると、識別符号の生成とは、特徴量空間を多数の小領域に分割して、それぞれの領域に対し識別符号を割りあてるものであると言える。このように空間を区切ることにより特徴量を識別する場合、ある程度の誤識別は避けられない。犯罪捜査という目的を考えると特徴量が類似している画像はある程度の誤識別を許容してでも広めに検索したい一方、プライバシー保護という目的を鑑みると個人識別符号は可能な限り誤識別を減少させたい。このトレードオフを調停するための基準として、人間の誤識別への主観的な許容度を導入することが考えられる。本研究では、顔特徴量抽出の手法を比較検討して各手法の誤識別の性能を測定する一方、誤識別に対する人間の主観的な許容度を被験者実験により計測する。

（２）検索正当性の確認手続きを備えた暗号化処理及び復号処理法前述の機構では、識別符号と暗号鍵及び復号鍵の対応付けの方法が最大の問題となる。顔画像の蓄積時に監視カメラが識別符号と鍵情報の対応付けを行うとした場合、対応付けについての情報を中立機関と通信する必要があるが、監視カメラが多数になると中立機関の負担は重くなり、また監視カメラは必ずしも通信が可能な環境にあるとは限らない。一方監視カメラは中立機関によりあらかじめ識別情報と対応づけられた鍵情報を用いる方法も考えられるが、その場合監視カメラは取り得る全ての識別情報についての鍵情報を自身で保有する必要があり、そもそも被写体となる人物が予め予想できない監視カメラ映像にあっては非現実的である。そこで本研究では、この問題を解決するために ID ベース暗号を導入する。ID ベース暗号は暗号化の実行者が e-mail アドレスなど公開された ID を用いた暗号化を行う一方、信頼できる第三者機関は暗号化実行者との通信なしに対応する復号鍵を ID 所有者に発行できる暗号化できる。ここで顔画像から生成した識別符号を ID とみなし、ID 所有者を警察と見なすことで、復号鍵の発行の可否を中立機関が検証・裁可できる枠組みが構築できる。ただし ID ベース暗号は比較的計算量が大きい処理を伴うため、画像の

蓄積及び検索機構に組み入れる場合には対象となる処理速度の速い共通鍵を中間鍵として用いるなどの工夫が必要となる。本研究ではそのような工夫の適用を前提とし、提案する機構が実応用に耐えうるスループットを出せることを明らかにする。

４．研究成果

本研究で識別符号量子化の閾値決定に用いる手法を以下に示す。まず、顔画像から抽出した顔特徴量はｄ次元ベクトルで表現されるものとする。ここでは顔特徴量の抽出技術として LBP(Local Binary Pattern)のような技術を想定している。抽出された特徴量の各次元の値は、閾値によって分割された複数の区間のいずれかに量子化される。

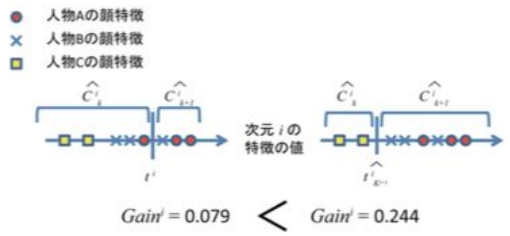

図１ 平均情報量を元にした閾値の決定

ここで用いる量子化の閾値は、既知の一般的な顔画像データベースから学習する。その閾値決定手続きは以下のようになる。顔特徴量のｉ次元目における区間を与える閾値を、t^i_k (k=1, ..., K^i -1)とする。K^iはｉ次元目の区間数である。更に、顔画像の顔画像全体集合をC, C^i_k(k=1, ..., K^i)をt^i_kによって分割された各区間に所属する顔画像の集合とする。このとき、与えられた画像集合中の人物ｘの生起確率 p(x;C^i_k)と、C^i_kの平均情報量 H(C^i_k)は以下の式により定義される。

$$p(x;\ C^i_k) = n^i_k(x)/N^i_k \quad (N^i_k = \Sigma_x\ n^i_k\ (x)) \cdots(1)$$

$$H(C^i_k) = -\Sigma_{(x\in I)}\ p\ (x;\ C^i_k)\ \log\ p\ (x;\ C^i_k) \cdots(2)$$

ただし､IはCに含まれる人物の集合を示す。次に、ｉ次元目の平均情報量の期待値は以下の式で求められる。

$$\Sigma^{Ki}_{(k=1)}\ H(C^i_k)N^i_k/\Sigma_k\ N^i_k \cdots(3)$$

そして、次の手順により各区間の閾値を決定する。

・最初の閾値は空集合に対応

・新しい閾値をｔとする。K^iをｔによってCを C'^i_k (k=1, ..., K^i=1)に分割した区間の数とする。まず、あｔを加えた後の平均情報量の期待値を計算する

$$\Sigma^{(Ki+1)}_{(k=1)}\ H(C'^i_k)\ N'^i_k/\Sigma_k\ N'^i_k \cdots(4)$$

N'^i_kは C'^i_kの顔画像の総数を示す。次に、式(3)と(4)から、平均情報利得を計算し、これを最大にするようなｔを新たな閾値の候補 $t'^i_{(Ki+1)}$とする(図 4)

$$Gain^i = \Sigma^{(Ki+1)}_{k=1}H(C'^i_k)\times N'^i_k/\Sigma_k N'^i_k - \Sigma^{(Ki+1)}_{(k=1)}H(C^i_k)\ \times N^i_k/\Sigma_k N^i_k \cdots(5)$$

・$t'^i_{(Ki+1)}$を各次元ｉについて計算し、$Gain^i$を最大化するような $t'^i_{(Ki+1)}$を新しい区間の閾値として選択する。そして、K^iを K^i+1 に更新する。

このような手順を繰り返すことによって、平均情報利得を最大化するような各区間の閾値を選択し、顔画像の特徴量量子化の閾値として採用する。このような量子化により得られた識別符号を元に、計算時間の計測による計算性能の評価と、被験者アンケートによる人間の一致度の判断と誤判定の許容度の分析を行うため、プライバシー保護顔画像検索システムを実装した。

表１ 顔画像の登録・参照にかかる平均時間

動作	平均時間 (msec)
登録	6.39
参照	4.48

まず、実装したシステムを用いて、顔画像の登録・参照に要した計算時間を計測した。対象とする画像は、縦 120pixel、横 96pixel、データ量が 11KB のモノクロ画像である。1 レコードあたりの登録・参照を 1000 回ずつ実施して計算の所要時間を計測した結果、一回あたりの各動作の平均時間は表１のようになった。登録時の計算性能について考えると、一般的な映像の記録に用いられる 30frame/sec のスループットが出れば十分と考えられ、これは１フレームあたりの処理速度が 33.3msec 程度に相当する。実験結果では、登録しに平均 6.39msec 程度の時間しか要していないため、少なくともこのデータ量の映像については実時間動作が可能であると言える。参照時の計算性能については、1 レコードあたり平均 4.48msec で参照できる。もし記録時に 30frame/sec で映像が記録されたと仮定すると、検索に要する時間の 33.3 / 4.48 = 7.43 倍の長さの映像を機械的に検索できることになり、目視による確認よりも高い効率の検索が実現できる。

プライバシー保護顔画像検索システムの機械的な判定結果の正誤については、FAR(False Acceptance Rate) ・ FRR(False Rejection Rate)という指標により評価することができる。FAR は他人受入率で、 他人を本人と誤って受け入れる確率を示す。FRR は本人拒否率で、 本人を誤って他人と判別する確率を示す。本手法においては、FAR は、検索クエリの人物と別人物の顔画像を検索結果として取得する割合を示す。この割合は、システムの安全性を示す指標となる。また FRR は、検索した人物と同人物の顔画像が取得されない、検索漏れの割合を示す。今回の実装では、LBP で抽出後主成分分析を適用し

て次元削減した次元数 d=200 の顔特徴量に対して、FAR=2%、FRR=50%の性能を出すシステムが動作している。ここで FAR に対して FRR が大きいのは以下の理由による。プライバシー侵害のリスクは FAR の大きさが主要因になるため FAR を可能な限り小さくしたい一方で、監視カメラにおいては同一人物が一定時間撮影されるため複数枚の顔画像が記録でき、FRR が大きい場合であっても検索が有効に働くと考えられることから、そのようなパラメータ設定を行っている。そのようにして得られた機械的な判定結果の画像を人間に示した場合、人間が正しく同一性を判断できるかどうか、また機械がもし判定を誤った場合にそれを受容できるかといったことを評価するためのアンケート調査を行った。このアンケートは公募により集まった大学生 30 人に対して実施した。大問 A、大問 B では、検索のクエリ画像と結果画像の対を見せ、人間が正しく同一性を判断できるかどうか、また機械がもし判定を誤った場合にそれを受容できるかについての質問を行った。画像対としては、機械的な判定結果が True Positive、False Negative、 False Positive、 True Negative であった例をそれぞれ 4 例ずつ、計 16 例を用意した。

表 2　大問 A のアンケート結果

判定	ラベル	pair	I	II	III	IV	V	中央値カテゴリ
TP	A	f1	14	14	1	1	0	II
		f2	23	5	0	2	0	I
	NA	f3	4	13	1	7	5	II
		f4	1	9	7	11	2	III
FN	A	f5	27	3	0	0	0	I
		f6	14	14	2	0	0	II
	NA	f7	10	15	3	1	1	II
		f8	4	18	3	5	0	II
FP	A	f9	3	6	3	10	8	IV
		f10	0	0	1	4	25	V
	NA	f11	0	0	1	4	25	V
		f12	0	0	3	4	23	V
TN	A	f13	0	2	2	14	16	V
		f14	0	0	2	10	18	V
	NA	f15	0	3	2	6	19	V
		f16	0	0	1	6	23	V

大問 A のアンケート集計結果は表 2 のようになった。回答のカテゴリはそれぞれ、I.同一人物に見える、II.どちらかというと同一人物に見える、III.どちらともいえない、IV.どちらかというと同一人物に見えない、V.同一人物に見えない、である。集計結果中の「判定」とは、画像が前述の True Positive、False Negative、 False Positive、 True Negative などの判定結果のどれに対応するかの頭文字を示す。集計結果中の「ラベル」とは、人間の判断と機械の判定が一致しない場合について注目するために、おなじ判定結果のグループの中でも比較的似ている画像対を「A(Alike)」、似ていない画像対を「NA(Not Alike)」として、アンケート実施者が目視によりラベル付けしたものである。

画像例では、TP と FN のグループが同一人物の顔画像対であり、FP と TN のグループが別人物の顔画像対である。集計結果を見ると、TP・FN のグループと FP・TN のグループの間で明確な回答の差があり、かつ人間の回答は実際の人物の同一性を反映している一方、TP と FN の間及び FP と TN の間では回答に明確な差が見られない。このことは、人間の判断は機械の判定と一致しないことが多い一方、人間は機械よりも同一性を良く判断できることを示すと考えられる。また、ラベルに注目した場合、f9 の画像が「機械は同一と誤判定したが、人間も同一と判断しがちである」例と言えるが、全体の傾向を覆すほど、人間が回答を誤った例とまでは言えない。従って、これらの集計結果からは人間よりも機械の方が一致判定に成功するような事例は観察できなかった。

表 3　大問 B のアンケート結果

判定	ラベル	pair	I	II	III	IV	V	中央値カテゴリ
TP	A	f1	9	5	9	6	1	III
		f2	9	9	4	7	1	II
	NA	f3	2	10	4	9	5	III
		f4	4	11	7	6	2	II & III
FN	A	f5	12	8	3	4	3	II
		f6	10	6	4	7	3	II
	NA	f7	9	11	3	5	2	II
		f8	7	11	4	7	1	II
FP	A	f9	3	8	4	10	5	III
		f10	0	2	5	13	10	IV
	NA	f11	0	2	4	9	15	IV & V
		f12	0	1	4	11	14	IV
TN	A	f13	0	2	7	17	4	IV
		f14	0	3	7	9	11	IV
	NA	f15	0	4	7	11	8	IV
		f16	1	2	2	9	16	V

大問 B のアンケート集計結果は表 3 のようになった。回答のカテゴリはそれぞれ、I.許容できる、II.どちらかというと許容できる、III.どちらともいえない、IV.どちらかというと許容できない、V.許容できない、である。集計結果を大問 2 と比較した場合、ほぼ同じ傾向を持つことが読みとれる。従って、誤判定を許容できるかどうかは人間が正しく同一性を判断できるかどうかの結果とほぼ同様になる一方、機械の判定が人間と一致している場合には実際の結果が誤っていても許容される可能性がある。現状では人間の方が機械的な判定よりも正確に人物の同一性を判断できることから、社会的に検索結果の誤判定を受容できるかどうかという課題には、結局一致判定性能の向上が重要な要因であるということが示されていると言える。

上記をまとめると、計算時間の計測では、提案システムが市場で入手可能な一般的な計算機上で実時間動作するという結果が得られた。また、アンケート結果によれば、人間の判断は機械の判定と一致しないことが

多い一方で、機械よりも人間の方が同一性を良く判断でき、また判定誤りの許容度についても実際の同一性の判断とほぼ同様であることが結果として示された。

＜引用文献＞
① 馬場口登、プライバシーを考慮した映像サーベイランス、情報処理、Vol.48、No.1、2007、pp.30-36、http://id.nii.ac.jp/1001/00066042/

５．主な発表論文等
（研究代表者、研究分担者及び連携研究者には下線）

〔学会発表〕（計２件）
① 森村吉貴、舩冨卓哉、上原哲太郎、美濃導彦、プライバシ保護顔画像検索システムの実装評価、電子情報通信学会マルチメディア情報ハイディング・エンリッチメント研究会、2015 年 03 月 12 日～2015 年 03 月 13 日　大濱信泉記念館（沖縄県・石垣市）
② 舩冨卓哉、川西康友、森村吉貴、美濃導彦、満上育久、プライバシを考慮した防犯カメラ映像処理、人工知能学会全国大会、2013 年 06 月 04 日～2013 年 06 月 07 日、富山国際会議場ほか（富山県・富山市）

６．研究組織
(1)研究代表者
森村　吉貴（MORIMURA, Yoshitaka）
京都大学・学術情報メディアセンター・助教
研究者番号：８０５７８２７９